Struct Chem (2010) 21:965–976
DOI 10.1007/s11224-010-9633-7

ORIGINAL RESEARCH

Molecular structure of basic oligomeric building units of heparan-sulfate glycosaminoglycans

Milan Remko · Piet Th. Van Duijnen · Ria Broer

Received: 17 March 2010 / Accepted: 1 June 2010 / Published online: 17 June 2010

Abstract This study reports in detail the results of systematic large-scale theoretical investigations of the acidic dimeric structural units (D–E, E–F, F–G, and G–H) and pentamer D–E–F–G–H (fondaparinux) of the glycosaminoglycan heparin, and their anionic forms. The geometries and energies of these oligomers have been computed using HF/6–31G(d), Becke3LYP/6–31G(d), and Becke3LYP/6–311+G(d,p) methods. The optimized geometries indicate that the most stable structure of these units in the neutral state is stabilized via a system of intramolecular hydrogen bonds. The equilibrium structure of these species changed appreciably upon dissociation. Water has a remarkable effect on the geometry of the anions studied. Because of high negative charge, the solvent effect also resulted in an appreciable energetic stabilization of biologically active anionic forms of these glycosaminoglycans. The stable energy conformations around glycosidic bonds found for dimers and pentamer investigated are compared and discussed with the available experimental X-ray structural data for the structurally related heparin-derived pentasaccharides in cocrystals with proteins.

Keywords Glycosaminoglycans · Dimers and pentamer · DFT · Molecular structure · Solvent effect

Introduction

Heparan sulfates are complex saccharides found on all cell surfaces, which bind selectively to a variety of proteins and pathogens and play important roles in numerous physiological and pathological processes, such as angiogenesis, cancer, hemostasis, growth factor control, anticoagulation, and cell adhesion [1–5]. Many of these activities have been discovered using heparin. Heparin is a member of the heparan sulfate family of glycans. It is a heterogeneous polymer constituted of repeating disaccharide sequences of hexuronic acid (D-glucuronic, L-iduronic) and D-glucosamine [6, 7]. These residues are substituted with *N*- or *O*-sulfate groups with various degrees of substitution. The chemical heterogeneity of heparin in terms of its sulfation pattern and backbone chemical structure facilitates binding to a variety of proteins such as growth factors, enzymes, and morphogenes [8]. Heparin is strongly acidic because of its content of covalently linked sulfate and carboxylic acid groups. This complex organic acid, which is found especially in lung and liver tissue, has a mucopolysaccharide as its active constituent, prevents platelet agglutination and blood clotting, and is therapeutically used in the form of its sodium salt in the treatment of thrombosis and in heart surgery. Its antithrombotic activity is explained by the interaction with protein antithrombin III (AT-III) [7, 9–11]. The action of heparin starts when it binds to antithrombin through a group of five subunits (DEFGH). It was found that this unique pentasaccharide fragment (PS) constitutes the minimal binding domain for AT-III. It contains five *O*-sulfate groups, three *N*-sulfate groups, and two carboxylate moieties. In the

Electronic supplementary material The online version of this article (doi:10.1007/s11224-010-9633-7) contains supplementary material, which is available to authorized users.

M. Remko (✉)
Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Comenius University Bratislava, Odbojarov 10, 832 32 Bratislava, Slovakia
e-mail: remko@fpharm.uniba.sk

P. Th. Van Duijnen · R. Broer
Department of Theoretical Chemistry, Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

antithrombin–pentasaccharide complex, these acidic groups are completely ionized and interact with the complementary amino acid side chains (Arg, Lys, Glu, and Asn) on the protein [11–15]. Despite its importance, the 3D molecular structure of native heparin and its simpler derivatives are not known. The absence of experimental structural data of basic building units of heparin presents a challenge to the application of molecular modeling methods in order to obtain insight into the recognition and binding processes.

The molecular structure of sodium salts of monomeric, dimeric, trimeric, and pentameric structural units of heparin was investigated using the density functional theory [16–21]. Hricovíni et al. examined structure and spin–spin coupling constants of monomeric and dimeric units of glucosaminoglycans [22–24]. Molecular structure of structurally related hyaluronan was recently reviewed [25], and conformational properties of the disaccharide-building units of hyaluronan were investigated using the density functional theory (DFT) [26]. Some of novel anticoagulants were also treated theoretically [27]. However, higher oligomeric structures of native glycosaminoglycans carrying free acidic sulfate and carboxyl groups have not been quantum chemically examined before. In this article, we have used large-scale quantum chemical calculations for the study of stable geometries of four disaccharide and pentasaccharide (DEFGH) units of heparin. Of particular interest are the molecular geometries and how these properties are changed upon dissociation and/or solvation. The results of this study are discussed and compared with the available experimental data for structurally related systems and discussed with the present theories of action of these glycosaminoglycans.

Computational details

Theoretical calculations of both, neutral acids and anions of disaccharides D–E, E–F, F–G, and G–H, and pentamer D–E–F–G–H of the heparin (Fig. 1) were carried out with the Gaussian 03 computer code [28] at the ab initio SCF (HF [29]) and (DFT [30], and Becke3LYP [31–34]) levels of theory using the 6–31G(d) and 6–311+G(d,p) basis sets. The structures of all gas-phase species were fully optimized at the HF/6–31G(d) and Becke3LYP/6–31G(d) levels of theory without any geometrical constraint. In order to check the correctness of the B3LYP-calculated geometries using the double-ζ basis set, we also performed calculations of the heparin species, using the basis set of the triple-ζ quality (B3LYP/6–311+G(d,p) level of theory) implemented in the Gaussian 03 package of computer codes [28]. The solvation energies of the neutral and anionic species in water solution were carried out using two solvation methods, CPCM [35, 36] and Onsager [37] models.

For identification of low energy conformations of the five oligomers studied, conformational analysis is necessary. A conformational search using theoretical methods is, for such large molecules, limited to the application of some of the available force-field methods. However, there is still no molecular mechanics force-field parametrization capable of adequately reproducing all glycosaminoglycan conformational features [9, 38]. It is common in the theoretical investigation of glycosaminoglycans to use structural data obtained from experimental NMR spectroscopy or X-ray crystallography as guides to the quality of the calculations.

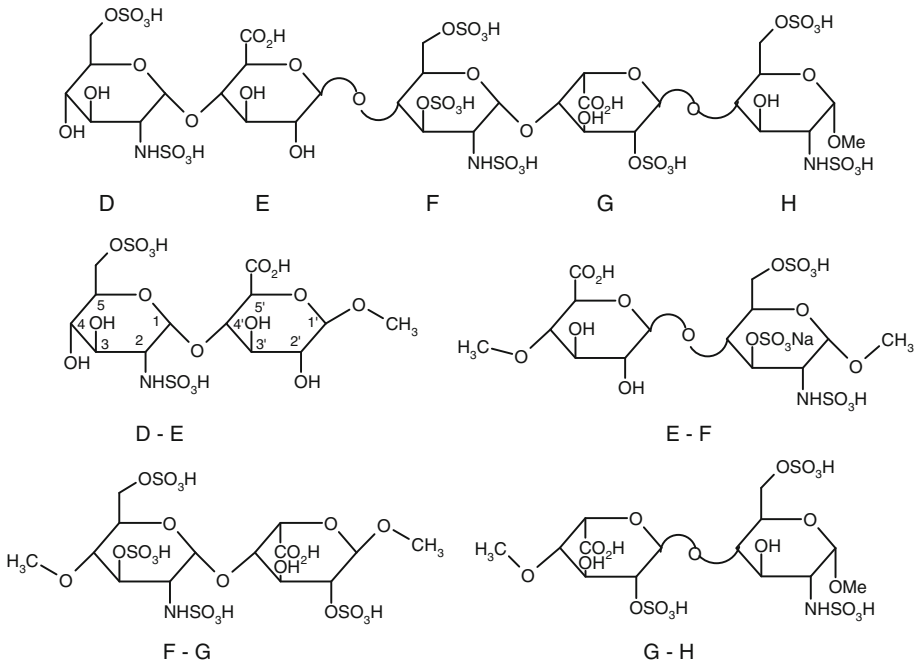

Fig. 1 Structure and atom numbering of the glycosaminoglycan species investigated

Initial structures for theoretical calculations were constructed by means of the Gauss View graphical interface of Gaussian. The orientation of the OH, COOH, $NHSO_3H$, and OSO_3H groups in gaseous phase was carried out in such a way that, where possible, the structures with the most probable intramolecular hydrogen bonds were assumed. As our previous calculations have shown, the most stable structures of the acidic monomeric structural units of heparin are stabilized via such intramolecular hydrogen bonds [16]. The pyranose rings of the D-glucosamine in these dimers were considered in more stable $^4C_1$ conformation. The starting conformation of the L-iduronic acid building unit G in dimers F–G and G–H was set to the more stable skew-boat ^2S_o form [4]. This conformation was taken, as it is the prevalent form of this residue in heparin and its model compounds [1, 9, 16, 18]. Initial conformations used in the DFT calculations of the dimers studied were based on the previous calculations of corresponding sodium salts [17]. The relative orientation of pyranose rings defined by two torsion angles (Φ, Ψ) at the (1 → 4) glycosidic linkage was adjusted according to the experimental data for structurally related heparin-derived oligosaccharides [39, 40].

Results and discussion

General considerations

It is well known that heparin is a long, linear chain sugar composed of repeating disaccharide units. The 24 disaccharides that can be formed from basic monosaccharides are the basic sequence units of heparin and heparan sulfate [4]. Of these, IndoA(2S)-(1 → 4)GlcNS(6S) (dimer G–H, Fig. 1) is the most abundant in heparin. Disaccharides composed from other combinations of the monosaccharides represent minor repeating dimeric units of heparin. The D–E, E–F, F–G, and G–H disaccharidic units of heparin investigated by us are part of the unique antithrombin-binding pentasaccharide sequence DEFGH in which two neighboring monomeric units of heparin bound by the (1–4) glycosidic bonds are substituted with the methyl groups (Fig. 1). The initial calculations were carried out using the HF/6–31G(d) method. Initial optimization at the lower level of theory shortened the overall time for the calculation and provided better starting structures for the higher-level calculations. Although we do not report detailed results for HF calculations in the gas state, we note that the equilibrium geometries computed at the HF level of theory are in general agreement with the DFT results obtained with the "standard" 6–31G(d) basis set. The superposition of the 3-D equilibrium structures of the pentasaccharide D–E–F–G–H manifesting the overall

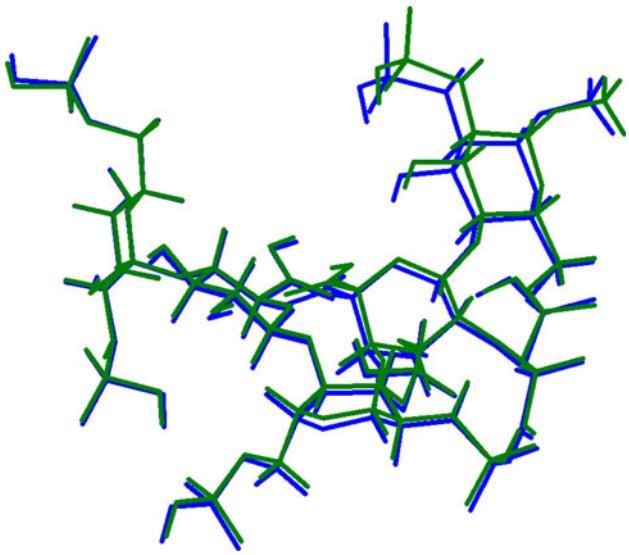

Fig. 2 Molecular superimposition of the gas-phase optimized structure of pentasaccharide DEFGH computed at the HF/6–31G(d) level of theory (*green*) and pentamer optimized using the B3LYP/6–3G(d) method (*blue*). (Color figure online)

difference in geometries of this oligomer computed by two quantum chemical methods is shown on Fig. 2. In order to study the basis set effects on the geometries of the glycosaminoglycans investigated within the DFT theory, we also carried out calculations of dimeric and pentameric structures using the larger basis set 6–311+G(d,p). The extension of the basis set in the DFT calculations resulted in only small changes in the equilibrium geometry of the systems studied. The optimized torsion angles at the glycosidic bonds using two basis sets within the DFT theory fit one another to within about 2°–5° (Tables 1, 2, 3). Thus, for purposes of determining reasonable geometries of larger oligomers of glycosaminoglycans, the B3LYP/6–31G(d) should be considered the preferred method of choice. The Cartesian coordinates (Å) of all gas-phase dimers investigated, optimized at the B3LYP/6–311+G(d,p) level of theory, are given in Table A of the electronic Supporting Information.

The effect of bulk solvent is treated with two solvation methods (the Onsager model [37] and CPCM [35, 36]) for comparison. Initial calculations were carried out, for computational reasons, using the SCRF formalism of Wong et al. [41]. The radii of the cavities used in this approach were chosen after a volume calculation of each molecule, and the dielectric constant of water ($\varepsilon = 78.39$) was used. The placing of the isolated molecules into a spherical cavity within a dielectric medium of the Onsager model of solvation does not represent the realistic situation in the biological medium; it seems helpful in revealing the main role of the solvent in intermolecular electrostatic interactions. The second, CPCM (conductor-like polarizable model), defines the cavity by the envelope of spheres

Table 1 Optimized glycosidic torsion angles (°) of the dimers studied

Dimer	Φ	Ψ	Φ, exp	Ψ, exp	Φ_H	Ψ_H	Φ_H, exp	Ψ_H, exp
D–E			89[a], 83[c]	−141[a], −158[c]			−49[b] (−42[c])	−31[b] (−40[c])
HF/6–31G(d)	76.7	−149.7			−39.7	−30.3		
B3LYP/6–31G(d)	74.5	−147.2			−41.4	−28.8		
B3LYP/6–311+G(d,p)	76.1	−150.4			−39.9	−31.6		
Solvated molecule[d]	77.4	−149.1			−38.9	−29.9		
Na salt[e]	59.7	−161.0						
E–F			−86[a], −75[c]	−113[a], −96[c]			43[b] (49[c])	12[b] (22[c])
HF/6–31G(d)	−67.0	−115.7			53.5	5.4		
B3LYP/6–31G(d)	−70.6	−115.9			50.1	4.5		
B3LYP/6–311+G(d,p)	−65.3	−113.6			55.3	7.4		
Solvated molecule[d]	−74.7	−114.3			46.2	5.7		
Na salt[e]	−74.5	−98.5						
F–G			55[a]	−155[a]				
HF/6–31G(d)	69.2	−163.8			−49.3	−43.9		
B3LYP/6–31G(d)	67.2	−159.3			−50.8	−39.7		
B3LYP/6–311+G(d,p)	70.0	−160.5			−47.6	−41.0		
Solvated molecule[d]	82.1	−141.2			−34.4	−21.0		
Na salt[e]	101.5	172.5						
G–H			−70[a]	−115[a]				
HF/6–31G(d)	−67.9	−113.2			50.5	8.5		
B3LYP/6–31G(d)	−72.6	−111.7			44.8	9.7		
B3LYP/6–311+G(d,p)	−64.6	−109.0			53.1	12.8		
Solvated molecule[d]	−77.4	−123.8			40.5	−4.0		
Na salt[e]	–57.4	−129.7						

[a] X-ray data of the pentasaccharide DEFGH bound to the antithrombin [12]

[b] NMR data for free heparin-derived hexasaccharide [48]

[c] X-ray data of the heparin-derived hexasaccharide bound to the fibroblast growth factor [49]

[d] B3LYP/6–31G(d) calculation

[e] B3LYP/6–31+G(d) structure taken from the reference [17]

centered on the atoms or the atomic groups [35, 36]. The whole concept of using such macroscopic properties as dielectric constants in microscopic computations has been criticized [42, 43]. Despite all these valid criticisms, continuum-based methods of solvation are used extensively and successfully in a variety of problems [44, 45]. It has been shown previously that the conductor-like polarizable method reproduces hydration energies with accuracy in the order of a few kcal/mol but mostly (70% of the cases) even better than 1 kcal/mol [46]. The solvent effect has a remarkable effect on the geometry of the glycosaminoglycans studied. The solvated phase energies within the Onsager model were obtained after full geometry optimization in water. However, glycosaminoglycans studied failed to optimize geometries in water within the CPCM formalism. Table 4 contains water stabilization energies using single-point CPCM calculations and in vacuo fully optimized geometries. The CPCM provided substantially more stable structures than Osanger's model, with the solvent stabilization being especially apparent in the case of anionic species. The CPCM-computed relative solvation energies correlate qualitatively well with the degree of ionization of individual dimeric and pentameric species (Table 4).

Dimeric and pentameric structural units

Dimer D–E

The optimized structure of dimer D–E is stabilized by means of system of intramolecular hydrogen bonds formed by hydroxyl groups and neighboring proton acceptors (Fig. 3). The conformation along the glycosidic linkage is stabilized via inter-ring hydrogen bond formed by the

Table 2 Optimized glycosidic torsion angles (°) of the anions of dimers studied

Anion	Φ	Ψ	Φ, exp	Ψ, exp	Φ_H	Ψ_H	Φ_H, exp	Ψ_H, exp
D–E			89[a], 83[c]	−141[a], −158[c]			–49[b] (−42[c])	–31[b] (−40[c])
HF/6–31G(d)	81.1	−156.0			−35.8	−38.0		
B3LYP/6–31G(d)	79.2	−155.5			−36.9	−38.1		
B3LYP/6–311+G(d,p)	84.1	−157.7			−32.3	−40.4		
Solvated anion[d]	84.4	–152.4			–32.8	–35.8		
E–F			−86[a], −75[c]	−113[a], −96[c]			43[b] (49[c])	12[b] (22[c])
HF/6–31G(d)	20.0	−122.9			139.3	−1.4		
B3LYP/6–31G(d)	27.1	−129.4			145.9	−8.8		
B3LYP/6–311+G(d,p)	23.5	−127.9			142.2	−7.0		
Solvated anion[d]	27.8	–129.8			146.1	–9.5		
F–G			55[a]	−155[a]				
HF/6–31G(d)	75.2	−149.8			−41.2	−29.6		
B3LYP/6–31G(d)	74.2	−148.6			−41.9	−28.5		
B3LYP/6–311+G(d,p)	75.3	−149.3			−40.7	−29.7		
Solvated anion[d]	73.9	–148.3			–42.3	–28.3		
G–H			−70[a]	−115[a]				
HF/6–31G(d)	−59.4	−94.6			59.1	29.4		
B3LYP/6–31G(d)	−53.2	−100.8			64.6	23.0		
B3LYP/6–311+G(d,p)	−58.5	−95.5			59.5	28.0		
Solvated anion[d]	–54.5	–110.8			62.5	12.3		

[a] X-ray data of the pentasaccharide DEFGH bound to the antithrombin [12]

[b] NMR data for free heparin-derived hexasaccharide [48]

[c] X-ray data of the heparin-derived hexasaccharide bound to the fibroblast growth factor [49]

[d] B3LYP/6–31G(d) calculation

Table 3 Optimized glycosidic torsion angles (°) of the pentamer and its anion studied

Angle	D–E–F–G–H				D–E–F–G–H anion		
	HF/6–31G(d)	B3LYP/6–31G(d)	Solvated molecule[a]	Na salt[b]	HF/6–31G(d)	B3LYP/6–31G(d)	Solvated anion[a]
Φ_1	74.7	76.7	77.4	63.7	68.6	66.0	76.6
Ψ_1	169.3	167.3	166.3	–169.8	–157.4	–157.7	–134.7
Φ_2	–105.4	–105.3	–106.0	–124.8	–93.9	–97.3	–100.3
Ψ_2	–92.8	–93.0	–94.5	–77.7	–75.3	–75.4	–76.6
Φ_3	70.9	72.4	69.2	54.3	78.1	80.9	76.9
Ψ_3	–148.4	–148.7	–148.9	–175.9	–163.7	–162.9	–166.5
Φ_4	–89.4	–93.3	–93.9	–69.1	–73.6	–73.0	–72.6
Ψ_4	–128.4	–131.0	–132.9	–127.1	–137.9	–140.4	–138.8
Φ_{H1}	–42.3	–40.2	–39.4	–54.4	–48.7	–50.6	–40.0
Ψ_{H1}	–71.8	–73.9	–75.0	–52.9	–39.5	–40.9	–17.3
Φ_{H2}	14.7	14.8	14.2	–5.3	30.2	27.1	23.9
Ψ_{H2}	24.8	23.9	22.7	42.2	48.1	47.9	46.2
Φ_{H3}	–46.9	–45.1	–48.2	–65.2	–39.7	–36.6	–40.8
Ψ_{H3}	–30.0	–30.7	–30.7	–57.7	–44.6	–44.1	−47.5
Φ_{H4}	31.5	26.9	26.4	49.4	45.4	45.3	45.7
Ψ_{H4}	–8.6	–12.2	–14.2	–11.2	–18.6	–21.1	–19.7

[a] B3LYP/6–31G(d) calculation

[b] B3LYP/6–31G(d) structure taken from the reference [20]

Table 4 The Becke3LYP/6–31G(d) solvent stability of the oligomers of heparin investigated[a]

No.	Oligomer	Cavity value[b] (a_0)	ΔE[c] Onsager	ΔE[c,d], CPCM	Gas-phase dipole moment[e]
1	D–E	5.94	−19.7	−192.8	9.17
2	E–F	5.73	−23.1	−220.6	9.47
3	F–G	6.13	−56.3	−242.0	6.52
4	G–H	5.93	−81.9	−220.0	11.71
5	D–E–F–G–H	7.91	−2.5	−356.6	4.45
6	D–E (anion)	5.52	−40.6	−1271.8	11.65
7	E–F (anion)	5.94	−21.3	−1995.7	9.89
8	F–G (anion)	6.10	−3.31	−2906.7	4.18
9	G–H (anion)	5.82	−49.5	−2012.3	13.03
	D–E–F–G–H (anion)	7.54	−7.9	−7021.4	8.08

[a] Water as solvent
[b] Ångström (Å)
[c] kJ/mol
[d] Single-point CPCM calculation
[e] Debye (D)

C(2)–$NHOSO_3H$ group of ring D and proton accepting C(3′)OH group of E moiety (Fig. 3). Water solvent does not have pronounced effect on the overall shape of this dimer (Table 1). The dissociation of three acidic hydrogen atoms leads to the structural change of the anionic sulfate and carboxylate groups. The space structure of a trianion of the dimer D–E is in the isolated state stabilized via intramolecular hydrogen bonds. All the four free hydroxyl groups of this dimer are involved in this system of intramolecular hydrogen bonds. The unsubstituted C(4)OH group takes part in the charged intramolecular C(4)O–H···$^{-}$O–S hydrogen bond. The strong electrostatic attraction in this hydrogen bond results in much shorter optimal hydrogen bond distance of 1.772 Å, which is appreciably less than the sum of the van der Waals radii [47] of hydrogen and oxygen atoms (2.7 Å). The overall 3-D shape of this anion is stabilized via charged H-bond formed by the C(3′)OH group and neighboring C(2)$NHSO_3^-$ group with an optimal O–H···$^{-}$O–S distance of 1.747 Å. Hydration of this anion caused appreciable geometry changes (Table 2), especially in the position of the bulky *N*-, and *O*-sulfated anionic groups. However, the hydrogen bond pattern observed in the gas-phase dimer D–E was also preserved in the solvated system.

Dimer E–F

Subunit F of this dimer is substituted by three sulfate groups and represents the most sulfated residue in heparin. The equilibrium positions of the C(2′)- and C(3′)-substituted sulfate groups in both neutral and anionic structures, respectively, is fixed by a system of H-bonds (Fig. 3). However, the ionization of sulfate groups is connected with large conformational rearrangement and a different stabilization of this dimeric structure in its anionic form (torsion angle Φ, Tables 1, 2). Hydration of this dimer caused conformational rearrangement of the functional sulfate groups and change of conformation around the glycosidic bond.

Dimer F–G

This dimer contains five polar acidic groups, which in the gas-phase structure are oriented out of the pyranose rings of this dimer. The conformation of two neighboring sulfate groups of unit F is stabilized by means of neutral intramolecular hydrogen bond of the SO–H···OS type (Fig. 3). The molecular structure of the pentanion in the gas-phase is stabilized via intramolecular hydrogen bond. The strong electrostatic attraction in the charged C(2)$NHOSO_3^-$···H–OC(3′) hydrogen bond with an optimal hydrogen bond distance 1.806 Å stabilizes conformation around the glycosidic bond. Appreciable structural changes of the F–G dimer upon solvent (water) are especially observed in the neutral species (Tables 1, 2).

Dimer G–H

The geometry of the optimized acid of the G–H dimer is stabilized by the system of intramolecular hydrogen bonds without interring coordination (Fig. 3). Similar conformation is found with the gas-phase anion. Every one of the pyranose rings of this dimer carries one free hydroxyl group, which is involved in intramolecular hydrogen bonds within individual monomers (Fig. 3). These hydrogen bonds are also preserved in the hydrated species.

Without the presence of the stabilization effect of interring coordination by means of hydrogen bonds, the solvent has a considerable effect on the conformational stability of this dimer around the glycosidic linkage (Tables 1, 2).

Pentamer D–E–F–G–H

The geometry of the optimized pentamer (fondaparinux) D–E–F–G–H is also presented in the Supporting Information.

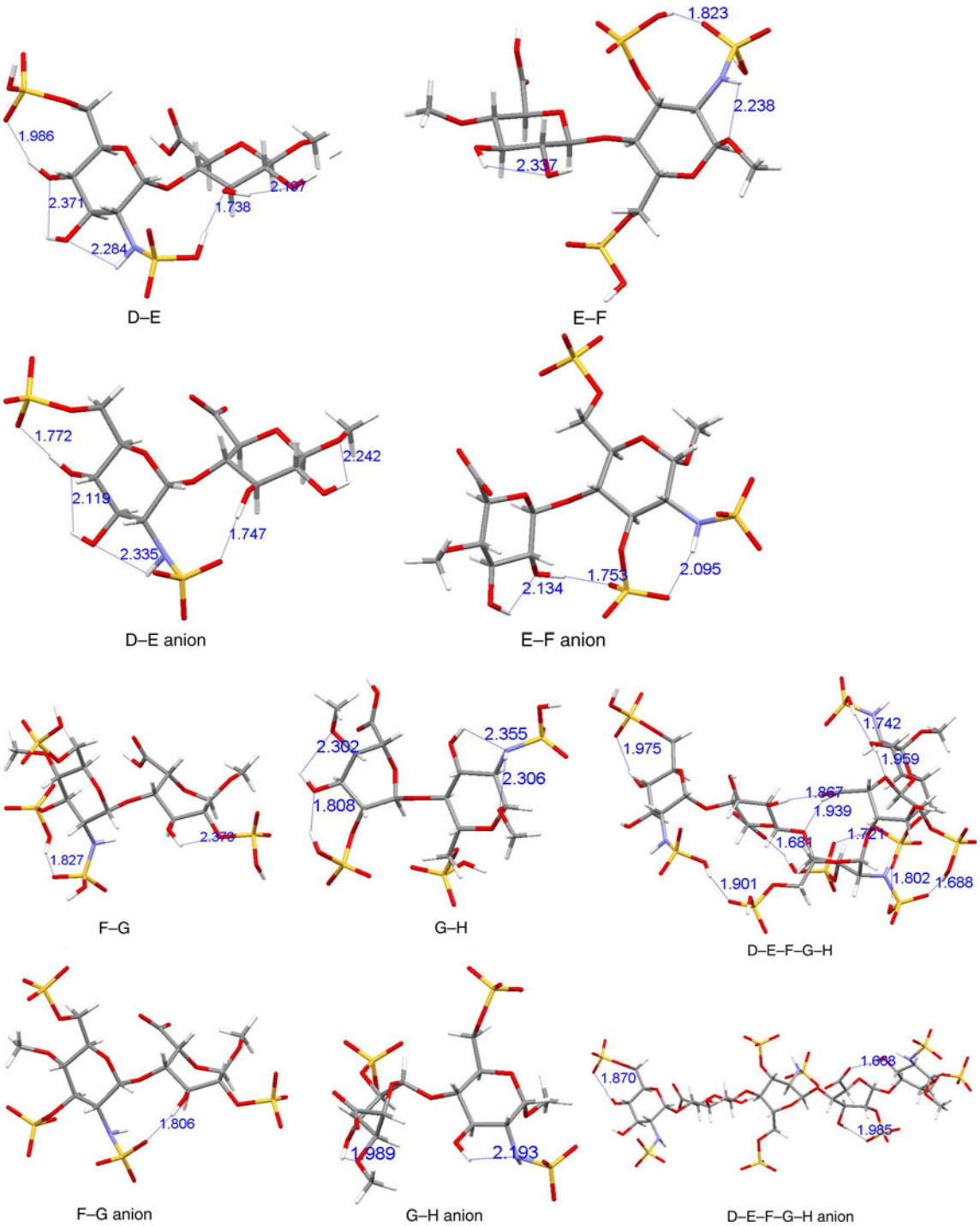

Fig. 3 B3LYP/6–31G(d) lowest energy structures of the disaccharide and pentasaccharide species studied. Interatomic distances are in Å

Ten acidic sulfate and carboxyl groups of fondaparinux are in the isolated molecule involved in a network of intramolecular hydrogen bonds stabilizing its pentasaccharide chain in a "bent" conformation. The solvent effect evaluated using the Onsager solvent reaction field does not change appreciably the equilibrium structure (Table 3). In the absence of the hydrogen bond stabilization conformation effect, the equilibrium conformation of the polyanion of pentasaccharide represents an extended structure (Fig. 3). The inter-ring conformation around the glycosidic bond in the G–H fragment is stabilized by charged hydrogen bond formed by the carboxylate group of the IndoA ring and C(3′)O–H group with an optimal hydrogen bond distance 1.668 Å. Aqueous solution modeled using the Onsager method does not have a

remarkable effect on the geometry of the pentamer species studied.

Structure of glycosidic linkage

For the definition of the torsion angles at the glycosidic bonds, we use the recommendations and symbols of nomenclature proposed by IUPAC (http://www.chem.qmul.ac.uk/iupac/2carb/). The relative orientation of a pair of dimers and pentamer studied is described by two torsion angles at the glycosidic linkage, denoted Φ and Ψ. For a (1 → 4) linkage the definitions of these torsion angles are as follows: $\Phi = O(5)–C(1)–O(1)–C(4')$ or $\Phi_H = H(1)–C(1)–O(1)–C(4')$; and $\Psi = C(1)–O(1)–C(4')–C(5')$ or $\Psi_H = C(1)–O(1)–C(4')–H(4')$. The glycosidic torsion angles of the fully optimized dimers and pentamer are given in Tables 1, 2, 3, 5. These tables also contain values of experimentally determined Φ and Ψ torsion angles for the complex of antithrombin with pentasaccharide [12], further available experimental data for glycosidic torsion angles in D–E and E–F dimeric units of free hexasaccharide measured by NMR spectroscopy [48], and in bound state in the complex of hexasaccharide with basic fibroblast growth factor protein obtained from X-ray diffraction studies [49]. In the absence of direct experimental X-ray structural data about the molecular structure of free acids of the species studied, the computed values for glycosidic torsion angles can be compared with the available data for structurally related oligomers only. Because of its highly acidic sulfate and carboxyl groups, heparin and its oligomeric species exist as anions at physiologic pH [6, 7, 16, 50]. Thus, the active form of heparinoid drugs is a polyanion with an average negative charge of individual saccharide residues of about 2.3.

The equilibrium structure of the four dimers investigated in the isolated state is determined by the coordination of the sulfate, carboxyl, and hydroxyl (hydroxymethyl) groups. This specific interaction gives rise to the unique overall shape of individual dimers (Fig. 3), and is manifested by the computed values of the Φ and Ψ torsion angles of the glycosidic linkage. Values of these glycosidic angles for individual dimers are considerably different (Table 1). Disaccharides investigated are part of the unique pentasaccharide-binding domain of heparin (also called the antithrombin III-binding domain), which is necessary to stimulate antithrombin activity [11]. The experimental values for the glycosidic angles in the pentasaccharide–antithrombin complex are well interpreted by the corresponding angles computed for isolated dimeric structural units (Table 1). The glycosidic angles for D–E and E–F linkages in the complexes of the pentasaccharide–antithrombin [12] and hexasaccharide–fibroblast growth factor [49] are close to each other, indicating that the heparin polymeric chain-binding sites by the interaction with different proteins do not undergo large conformational change (Table 1). The computed glycosidic angles for the D–E and E–F dimers are in good agreement with experimentally (NMR spectroscopy) determined Φ and Ψ torsion angles in structurally related hexasaccharide [48]. The X-ray data for the same glycosidic linkages in the complex of this hexasaccharide with the fibroblast growth factor [49] are slightly different (Table 1). According to our calculations, the inter-ring orientation of the both, isolated and hydrated anionic species of dimers studied is quite different. Experimental investigations indicate that heparin saccharides show relatively conserved angles Φ and Ψ in glycosidic linkages [9]. Accordingly, dimeric saccharide building units of heparin cannot represent all characteristic structural features of the native preparations. For obtaining a more detailed picture of the subtle structural features of the pentasaccharide-binding site of heparin, it is necessary to investigate higher oligomers of heparin using sophisticated quantum chemical methods.

The glycosidic torsion angles of the fully optimized fondaparinux are given in Table 3. The equilibrium structure of the pentamer investigated in the isolated state is determined by the multidentate coordination of the proton donor and proton acceptor groups of the sulfate, carboxyl, and hydroxyl (hydroxymethyl) moieties. This specific interaction gives rise to the unique overall shape of individual oligomers (Fig. 3), and is manifested by the computed values of the Φ and Ψ torsion angles of the glycosidic linkage (Tables 3, 5). The overall structure of the gas-phase anion is a compromise between the repulsion forces of anions and H bonding pattern in pentamer. Our calculations of D–E, E–F, F–G, and G–H dimers (Tables 1, 2) reached considerably different individual glycosidic torsion angles. Thus, the 3D structure of the pentamer of heparin cannot be deduced from the known conformation of the basic dimeric units. The same is true for the D–E–F and F–G–H trimeric units of heparin [20].

As regards of the fondaparinux (pentamer D–E–F–G–H), the dissociation of the 10 acidic groups is connected with the large change of the overall 3D structure. For both gas-phase and hydrated polyanion, the "extended" structure of the uniform helical conformation is the most stable one. The calculated helical parameter *h* (the axial rise per pentasaccharide) was as follows: 13.79 Å with solvated fondaparinux, and 22.42 Å in solvated polyanion of the fondaparinux. Inter-ring orientations in this pentamer are significantly influenced by the system of intramolecular hydrogen bonds, ionization, and/or solvent effect. The computed structural data can be directly compared with the gas-phase data for isolated systems and/or experimental data for molecules measured in water solution. However, to our knowledge, no experimental 3D structural parameters for dimers and pentamer studied are available till now. In the absence of experimental (gas-phase) data, the geometry

Table 5 B3LYP/6–31G(d) optimized glycosidic torsion angles (°) of the fondaparinux, its anion, and experimental structural data of related pentasaccharides in their complexes with proteins

Angle	D–E–F–G–H (fondaparinux), calculation				D–E–F–G–H, experiment[a]			
	DEFGH molecule	Solvated DEFGH molecule	Anion	Solvated anion	1AZX[b]	1E03[b]	1NQ9[c]	2GD4[d]
Φ_1	76.7	77.4	66.0	76.6	89.3 (91.0)	84.7 (96.9)	81.5 (77.7)	101.4
Ψ_1	167.3	166.3	–157.7	–134.7	−141.4 (−154.1)	−137.0 (−153.7)	−137.6 (−142.6)	−157.6
Φ_2	–105.3	–106.0	–97.3	–100.3	−86.0 (−79.6)	−87.3 (−72.4)	−89.0 (−86.5)	−84.4
Ψ_2	–93.0	–94.5	–75.4	–76.6	−112.7 (−118.8)	−112.5 (−116.8)	−108.6 (−107.4)	−120.8
Φ_3	72.4	69.2	80.9	76.9	55.4 (49.1)	58.7 (45.7)	59.2 (58.4)	63.1
Ψ_3	–148.7	–148.9	–162.9	–166.5	−154.4 (−153.2)	−164.1 (−151.1)	−162.1 (−159.5)	−157.0
Φ_4	–93.3	–93.9	–73.0	–72.6	−69.8 (−74.0)	−63.4 (−78.1)	−62.1 (−69.8)	−67.8
Ψ_4	–131.0	–132.9	–140.4	–138.8	−115.2 (−103.8)	−119.6 (−92.8)	−121.4 (−119.8)	−108.5

[a] The X-ray structure of the second molecule in the asymmetric unit is in the *parenthesis*, [b] values of experimentally determined Φ and Ψ torsion angles for the complex of antithrombin with pentasaccharide [12], and [c] X-ray experimental data for antithrombin in the pentasaccharide-bound intermediate state [51], [d] pentamer (fondaparinux) anion from the cocrystal with antithrombin and coagulation factor Xa [52]

of parent molecules can be compared only with X-ray data on structurally related pentasaccharides in their complexes with proteins (Table 5), namely, natural pentasaccharide (in the complex of antithrombin with pentasaccharide [12], and X-ray experimental data for antithrombin in the pentasaccharide-bound intermediate state [51]) and the synthetic version of the naturally occurring pentasaccharide fondaparinux (in the complex with antithrombin and coagulation factor Xa [52]). The values of the glycosidic angles of these pentasaccharides are also listed in Table 5.

In the solid state, the pentasaccharide DEFGH is bound to the protein in the form of an anion. The solid-state structure of bound pentasaccharide can be affected by the so-called packing effects, which can distort the structure. The published cocrystal structures of pentasaccharide with proteins have highlighted the ionic interactions between specific anionic sulfate and carboxyl moieties of pentasaccharide with the basic cationic amino acid residues of binding site on the protein [8, 12, 51, 52]. However, ionic contacts only are not sufficient to explain the optimal structural fit of pentasaccharide to the binding site of protein. It is assumed that binding of the pentasaccharide to the protein would also induce local distorsions in an otherwise uniform helical structure of pentasaccharide, which are manifested as changes in the glycosidic torsion angles. The naturally related pentasaccharide conformations in cocrystals with different proteins and/or environments (pdb files 1AZX, 1E03, 1NQ9) represented by the glycosidic torsion angles are close to each other (Table 5). However, the experimentally determined torsion angles at the glycosidic linkage of the structurally related pentasaccharide fondaparinux in its complex with antithrombin and coagulation factor Xa (pdb file 2GD4) are different (Table 5), indicating that the fine differences in the ionic interactions resulting from the different patterns of the sulfate, carboxyl, and methoxy substituents of these two pentasaccharides also contribute to the optimal structural fit of these compounds to the binding site of the protein.

Experimentally determined 3D structure of the fondaparinux studied corresponds to the bound molecule at the protein (pdb file 2GD4); therefore, the general structural motifs of fondaparinux can be compared with results for isolated molecule from theoretical methods only. The experimental values for the glycosidic angles in the fondaparinux–protein complexes are well interpreted by the corresponding angles computed for solvated anions of the fondaparinux. Despite the fact that some of the torsion angles computed by us for the anion of fondaparinux (e.g., angles Ψ_1, Φ_2, Ψ_3, and Φ_4) proved to be close to the experimentally determined torsion angles in cocrystal of this anion (Table 5), the overall structure of the bound and unbound fondaparinux is very different (Fig. 4). The experimental helical parameter *h* (the axial rise per pentasaccharide) is in the cocrystal of the fondaparinux with

Fig. 4 Molecular superimposition of the in-solution-optimized anion of the pentamer DEFGH (*blue*) and pentamer anion from the co-crystal with antithrombin and coagulation factor Xa, PDB 2GD4 (*violet*). (Color figure online)

antithrombin and coagulation factor Xa equal to 21.51 Å, which is close to the analogous value computed for the solvated anion of the fondaparinux (22.42 Å). However, with respect to the conformational flexibility of polysaccharide chain, it is possible to build more than one helix with a given set of helical parameters. Thus, the equilibrium conformation of the pentasaccharide bound to the protein occurs between two extreme positions, namely, the thermodynamically stable conformation of the neutral pentasaccharide and its anion. The computed torsion glycosidic angles (Φ_1, Ψ_1) and (Φ_2, Ψ_2) for the rotation of the D–E and E–F units in the pentasaccharide studied (Table 1) fit well to available experimentally (NMR spectroscopy) determined glycosidic angles (−49, −31) and (43, 12), respectively, in structurally related hexasaccharide [48].

Effect of sodium salt formation on the molecular structure

Heparin and its low-molecular weight derivatives are usually administered as the sodium salt. It is thus instructive to determine the effect of sodium salts on the overall structure of individual oligomeric units of heparin. Our previous extensive studies of molecular structure of sodium and lithium salts of monomeric [16, 18, 19, 21] structural units, sodium salts of dimeric [17], and trimeric, and pentameric [20] structural units of heparin have shown that, in neutral salts, ionic interactions are dominant as the negatively charged sulfate and carboxyl groups form multidentate coordination bonds with Na^+ and/or Li^+ cations. The equilibrium structure of the sodium salts of the D–E, E–F, F–G, and G–H dimers in the isolated state is stabilized trough the coordination of sodium cations with the negatively charged oxygen atoms of the sulfate, carboxyl, and hydroxyl groups. The result of these specific interactions is the unique overall shape of individual dimers, as it is manifested by the computed values of the Φ and Ψ torsion angles of the glycosidic linkage [17]. Interring conformation in the pentamer is appreciably stabilized by the system of coordination bonds formed by 10 Na^+ ions and functional groups of hexose rings [20].

Previous calculations have shown that the dissociation of the sodium cations from their coordination sites (carboxylate and *O*-sulfate groups) is thermodynamically much more favorable than the deprotonation of the corresponding acids of the 1,4-DiOMe IdoA2S [16]. Computations that include the solvent effect have shown that, in water, the relative stability of $Na^+\cdots{}^-$O-glycosaminoglycan ionic bonds is rapidly diminished, and both contact ion pairs and solvent-separated ion pairs might coexist [19]. The favorable thermodynamic and dissolution characteristics of sodium salts of heparin drugs are probably the main reason for their preferred use in clinical practice [1–4]. The glycosidic torsion angles computed for neutral acids and sodium salts of dimeric and pentameric structures studied are quite different (Tables 1, 3). The stable conformation of the sodium salt of pentamer is stabilized by means of multicoordination of sodium cations with negatively charged anionic groups of side-chain carboxyl and sulfate moieties. In addition, this structure is also stabilized by means of intramolecular hydrogen bonds originating from the hydroxyl groups of individual monomeric structural units of heparin [20]. Native heparin structural units, consisting of hexose rings, are stabilized via the intramolecular hydrogen bonding system only. Distinctive bonding characteristics of H^+ and Na^+ ions are responsible for profoundly different 3D packing of these two compounds (Fig. 5). As it is apparent from experimental and calculated pK_a values of sulfate and carboxyl acidic groups of essential structural units of heparin, both the sulfate and carboxyl groups are at physiological pH = 7.3 fully ionized [16, 53]. Sodium salts of heparin and its derivatives, as salts of weak acids, are almost completely ionized in solution. The conjugate bases of oligomers of heparin resulting from the dissociation of native acidic forms and/or salts in water solution undergo

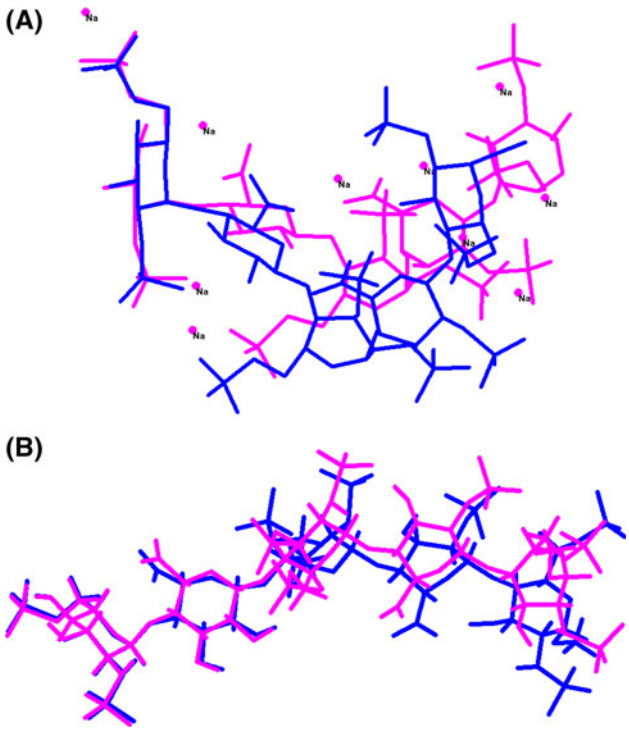

Fig. 5 **a** Molecular superimposition of the solution-phase-optimized structure of pentasaccharide DEFGH computed at the B3LYP/6–3G(d) level of theory (*blue*) and sodium salt of this pentamer optimized using the B3LYP/6–3G(d) method (*violet*). **b** Molecular superimposition of the solution-phase-optimized structure of (–10) anion of pentasaccharide DEFGH computed at the B3LYP/6–3G(d) level of theory (*blue*) and anion of sodium salt of this pentamer optimized using the B3LYP/6–3G(d) method (*violet*). (Color figure online)

large conformational changes. However, in water optimized molecular structures of two anions arising from the dissociation of free acid and/or sodium salt of pentasaccharide are very similar (Fig. 5).

Summary and conclusions

Quantum chemical calculations at the ab initio SCF and density functional levels of theory have been performed for conformations of dimers (D–E, E–F, F–G, and G–H) and pentamer D–E–F–G–H (fondaparinux) of the glycosaminoglycan heparin, and their ionized forms. For these oligomers of heparin, experimental structural data that consider their chemical and biological importance are rare. The following conclusions can be drawn.

(1) The optimized geometries indicate that the most stable structure of these structural units in the neutral state is stabilized via a system of intramolecular hydrogen bonds. The equilibrium structure of these species changed appreciably upon dissociation. Water has a remarkable effect on the geometry of the anions studied. Because of the high negative charge, the solvent effect also resulted in an appreciable energetic stabilization of biologically active anionic forms of these glycosaminoglycans.

(2) The equilibrium geometries computed at the HF level of theory are in good agreement with the DFT Becke3LYP results obtained with the "standard" 6–31G(d) basis set. The extension of the basis set to the triple-ζ quality in the DFT calculations resulted in only small changes in the equilibrium geometry of the systems studied. The optimized torsion angles at the glycosidic bonds using two basis sets within the DFT theory fit one another to within about 2°–5°.

(3) The experimental values for the glycosidic angles in the fondaparinux–protein complexes are well interpreted by the corresponding angles computed for solvated anions of the fondaparinux. However, the overall structure of the bound and unbound fondaparinux is very different.

(4) The computed structure of the gas phase of fondaparinux anion is a compromise between the repulsion forces of anions and H bonding pattern in pentamer. Our calculations of D–E, E–F, F–G, and G–H dimers have given considerably different individual glycosidic torsion angles. Thus, the 3D structure of the pentamer of heparin cannot be deduced from the known conformations of the basic dimeric units.

This study yields quantities that may be inaccessible or complementary to experiments and represents the first quantum chemical approach in which both the gas-phase and solvated phase 3D structure of the dimeric and pentameric units of native heparin are determined.

Acknowledgment The European Union HPC-Europa Transnational Access Program under the Project HPC-Europa2 (Project No 144) at SARA Amsterdam (M.R.) has supported this study. The study was also supported by Slovak Grant Agency VEGA contract No 1/0084/10 (M.R.). The authors acknowledge with thanks the Stichting Academisch Rekencentrum Amsterdam (SARA) for the use of its resources and for excellent support. M.R. thanks the Zernike Institute for Advanced Materials, University of Groningen, for its hospitality during his study stay in Groningen.

References

1. Desai UR (2004) Med Res Rev 24:151
2. Linhardt RJ (2003) J Med Chem 46:2551
3. Liu J, Thorp SC (2002) Med Res Rev 22:1
4. Rabenstein DL (2002) Nat Prod Rep 19:312
5. Sasisekharan R, Shriver Z, Venkataraman G, Narayanasami U (2002) Nat Rev Cancer 2:521
6. Lane DA, Lindahl U (eds) (1989) Heparin—chemical and biological properties. CRC Press, Boca Raton, FL
7. Casu B, Lindahl U (2001) Adv Carbohydr Chem Biochem 57:159
8. Raman R, Sasisekharan V, Sasisekharan R (2005) Chem Biol 12:267
9. Capila I, Linhardt RJ (2002) Angew Chem Int Ed 41:390
10. Conrad HE (1998) Heparin-binding proteins. Academic Press, San Diego
11. Petitou M, van Boeckel CAA (2004) Angew Chem Int Ed 43:3118
12. Jin L, Abrahams JP, Skinner R, Petitou M, Pike RN, Carrell RW (1997) Proc Natl Acad Sci USA 94:14683
13. van Boeckel CAA, Grootenhuis PDJ, Visser A (1994) Nat Struct Biol 1:423
14. Noti C, Seeberger PH (2005) Chem Biol 12:731
15. Huang L, Kerns RJ (2006) Bioorg Med Chem 14:2300
16. Remko M, von der Lieth WC (2006) J Chem Inf Model 46:1194
17. Remko M, von der Lieth WCJ (2006) J Chem Inf Model 46:1687
18. Remko M, Swart M, Bickelhaupt FM (2007) J Phys Chem B 111:2313
19. Remko M, van Duijnen PT, von der Lieth WC (2007) J Mol Struct Theochem 814:119
20. Remko M, von der Lieth WC (2007) J Phys Chem A 111:13484
21. Remko M, Hricovíni M (2007) Struct Chem 18:537
22. Hricovíni M, Bízik F (2007) Carbohydr Res 342:779
23. Hricovíni M (2007) Carbohydr Res 341:2575
24. Hricovíni M, Scholtzová E, Bízik F (2007) Carbohydr Res 342:1350
25. Hargittai I, Hargittai M (2008) Struct Chem 19:697
26. Pogány P, Kovács A (2009) Carbohydr Res 344:1745
27. Remko M (2009) J Mol Struct Theochem 916:76
28. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery Jr JA, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada

M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian 03, revision D.01. Gaussian, Inc., Wallingford, CT
29. Hehre WJ, Radom L, Schleyer PvR, Pople JA (1986) Ab initio molecular orbital theory. Wiley, New York
30. Bickelhaupt FM, Baerends EJ (2000) In: Lipkowitz KB, Boyd DB (eds) Reviews in computational chemistry, vol 15. Wiley-VCH, New York, p 1
31. Kohn W, Sham LJ (1965) Phys Rev A 140:1133
32. Becke AD (1988) Phys Rev A38:3098
33. Becke AD (1993) J Chem Phys 98:5648
34. Lee C, Yang W, Parr RG (1988) Phys Rev B37:785
35. Klamt A, Schűűman G (1993) J Chem Soc Perkin Trans II:799
36. Cossi M, Rega N, Scalmani G, Barone V (2003) J Compd Chem 24:669
37. Onsager L (1936) J Am Chem Soc 58:1486
38. Dyekjær JD, Rasmussen K (2003) Mini-Rev Med Chem 3:713
39. Ragazzi M, Ferro DR, Perly B, Sinaÿ P, Petitou M, Choay J (1990) Carbohydr Res 195:169
40. Imberty A, Lortat-Jacob H, Pérez S (2007) Carbohydr Res 342:430
41. Wong MW, Frisch MJ, Wiberg KB (1991) J Am Chem Soc 113:4776
42. van Duijnen PT, De Vries AH (1995) Int J Quant Chem Quant Biol Symp 29:523
43. Schultz CN, Warshel A (2001) Proteins 44:400
44. Tomasi J, Mennucci B, Cammi R (2005) Chem Rev 105:2999
45. Kongsted J, Söderhjelm P, Ryde U (2009) J Comput Aided Mol Des 23:395
46. Pye CC, Ziegler T (1999) Theor Chem Acc 101:396
47. Bondi A (1964) J Phys Chem 68:441
48. Mikhailov D, Linhardt RJ, Mayo KH (1997) Biochem J 328:51
49. Faham S, Hileman RE, Fromm JR, Linhardt RJ, Rees DC (1996) Science 271:1116
50. Gandhi NS, Mancera RL (2008) Chem Biol Drug Des 72:455
51. Johnson DJD, Huntington JA (2003) Biochemistry 42:8712
52. Johnson DJD, Li W, Adams TE, Huntington JA (2006) EMBO J 25:2029
53. Wang H, Loganathan D, Linhardt RJ (1991) Biochem J 278:689